

Lectures on Complexity and Black Holes

Lecture I

Leonard Susskind

Stanford Institute for Theoretical Physics and Department of Physics,
Stanford University, Stanford, CA 94305-4060, USA

Abstract

This is the first of three lectures on complexity and its role in black hole physics.

Contents

Preface

Lecture I: Hilbert Space is Huge

These lectures are a tale of two metrics on the same space—the space of states of a quantum system of K qubits. One metric is familiar to you and the other probably very unfamiliar. They are extremely different measures of the distance between quantum states and they have different purposes.

The two metrics are totally dissimilar. One, the inner product metric, is small in the sense that no two points are more distant than $\pi/2$. The other, relative complexity, is huge; almost all points are separated by a distance exponential in K . The inner product metric is positively curved like a sphere. The geometry of relative complexity is negatively curved, with the radius of curvature being much smaller than the maximum distance between points (the diameter of the space). Both are compact and homogeneous.

Although they both represent the distance between quantum states they are in a sense incommensurate: The distance between two points can be very small in one metric and exponentially large in the other. They represent very different relations between states.

Relative complexity should be of interest in quantum-computer science but has not been of much use in the study of ordinary quantum systems, not even black holes if we are interested in the region outside the horizon. It is only when we ask about the interior of a black hole, the region behind the horizon, that relative complexity takes its place as a fundamental quantity.

You might ask why put so much effort into understanding what can never be seen. The answer is two-fold: First Einstein's general theory of relativity predicts and describes the region behind the horizon and if we want to understand quantum gravity we have to take account of this fact. We will not have understood how the full geometry of GR emerges from quantum mechanics without following space into the black hole interior.

The second point is that in a cosmology like ours—de Sitter like—everyone is behind someone else's horizon. We won't understand anything without understanding horizons.

1 How Huge?

Let's consider the space of states of K qubits and make a simple estimate of its size. By size I don't mean the dimensionality of the space defined by the number of mutually orthogonal vectors. I mean something more like the total number of unit vectors. It's of course infinite but we can regulate it.

The dimension of the Hilbert space is 2^K and a general vector has the form,

$$|\Psi\rangle = \sum_1^{2^K} \alpha_i |i\rangle$$

The α are arbitrary complex numbers. At the moment we won't worry about normalizing $|\Psi\rangle$ or dividing out the overall phase. Now let's regulate the infinities by restricting each α_i to be one of m values. The total number of states is,

$$\#states = m^{2^K} = \exp(2^K \log m) \quad (1.1)$$

For $K = 4$, $m = 4$ the number of states is 4,294,967,296.

The logarithm of the number of states is more manageable,

$$\log(\#states) = 2^K \log m \approx 22. \quad (1.2)$$

There are two interesting things about ???. The first is how strongly it depends on K , namely it grows as 2^K . The second is how weakly it depends on the regulator parameter m . We'll see this trend many times in what follows.

2 Volume of CP(N)

To do any kind of rigorous counting of points in a continuous space we have to coarse-grain the space. An example is counting states in classical statistical mechanics where we coarse grain, replacing points by little balls of radius ϵ . This allows us to count states, for example to define entropy,

$$S = \log(\#states) \quad (2.3)$$

at the cost an additive term, $\Delta \log \epsilon$, Δ being the dimension of the phase space. In order to count states in $CP(2^K - 1)$ or unitary operators in $SU(2^K)$ we have to do something similar.

The space of normalized states with phase modded out is the projective space $CP(2^K - 1)$. Let's calculate its volume using the usual Fubini-Study metric. The answer is not hard to guess: it is the volume of the sphere of the same dimension divided by 2π . The volume of $CP(N)$ is,

$$V[CP(N)] = \frac{\pi^N}{N!} \approx \left(\frac{e\pi}{N}\right)^N \quad (2.4)$$

As I said, in order to count states we need to regulate the geometry. The simplest way to do that is to replace points by small balls of radius ϵ . The volume of epsilon-ball of dimension $2N$ is

$$V[B(2N)] = \frac{\pi^N}{N!} \epsilon^{2N} \approx \left(\frac{e\pi}{N}\right)^N \epsilon^{2N}. \quad (2.5)$$

The obvious thing to do is to divide the volume of $CP(N)$ by the volume of an epsilon-ball. The result is very simple. The number of epsilon-balls in $CP(N)$ is ϵ^{2N} .

Next we replace N by $2^K - 1$. We identify the number of states with the number of epsilon-balls in $CP(2^K)$.

$$\#states = \left(\frac{1}{\epsilon}\right)^{2(2^K-1)} \quad (2.6)$$

or

$$\log(\#states) = 2(2^K - 1) \log(1/\epsilon) \quad (2.7)$$

For $K = 4$ and $\epsilon = 1/2$ this gives 1,073,741,824 ϵ -balls.

Equation ?? shows the same pattern as ??. It grows rapidly like 2^K , and is only logarithmically sensitive to the cutoff parameter.

3 Relative Complexity

The enormity of the space of states of quantum systems begs the question of how to measure the distance between states. Are some states close to each other and others

enormously far? The usual metric in state space (by which I mean $CP(N)$) is the inner-product metric. It's also called the Fubini-Study (FS) metric. The FS distance between two vectors $|A\rangle$ and $|B\rangle$ is,

$$d_{fs}(A, B) = \arccos |\langle A|B\rangle|. \quad (3.8)$$

It varies from $d_{fs} = 0$ when $|A\rangle = |B\rangle$ to $\pi/2$ when $|A\rangle$ and $|B\rangle$ are orthogonal. The space with the FS metric is “small” in the sense that the furthest distance between points is $\pi/2$.

The FS metric has its purposes but it fails to capture something important. Imagine an intricate piece of machinery with a microscopic switch—a control qubit—that can be flipped on or off. Compare the distance between these two states,

$$|\mathbf{on}\rangle \quad |\mathbf{off}\rangle$$

with the distance between two other states: a brand new machine, and a pile of rust,

$$|\mathbf{new}\rangle \quad |\mathbf{rust\ pile}\rangle$$

The FS distance between states $|\mathbf{new}\rangle$ and $|\mathbf{rust\ pile}\rangle$ is exactly the same as the distance between $|\mathbf{on}\rangle$ and $|\mathbf{off}\rangle$, namely $\pi/2$. If the only criterion was the FS distance it would be just as easy to:

1. Make a transition from $|\mathbf{new}\rangle$ to $|\mathbf{rust\ pile}\rangle$ as to make a transition from $|\mathbf{on}\rangle$ to $|\mathbf{off}\rangle$.
2. Create a coherent superposition of $|\mathbf{new}\rangle$ and $|\mathbf{rust\ pile}\rangle$, or a coherent superposition of $|\mathbf{on}\rangle$ and $|\mathbf{off}\rangle$.
3. Do a measurement that would be sensitive to the relative phases of such superpositions in the two cases.

But of course this is not true. It is much easier make a superposition of $|\mathbf{on}\rangle$ and $|\mathbf{off}\rangle$ (it just involves superposing two single-qubit states) than of $|\mathbf{new}\rangle$ and $|\mathbf{rust\ pile}\rangle$. Evidently the FS metric fails to capture the difference between these two cases.

If it were equally possible to physically apply *any* operator to a system, then it would be no harder to superpose the machine and the rust pile, than the on-off states. But in the real world an operator that takes a machine to a rust pile is much more *complex* than an

operator that flips a qubit. In practice, and maybe in principle, we carry out very complex operations as sequences of simple operations. Once a set of “allowable” simple operations has been identified, a new measure of distance between states becomes possible. It can be defined as:

The minimal number of simple operations needed to go from $|A\rangle$ to $|B\rangle$.

In (quantum)information theory the simple operations are called gates and new distance measure is called relative complexity:

The relative complexity of $|A\rangle$ and $|B\rangle$ is the minimum number of gates required to go from $|A\rangle$ to $|B\rangle$.

One might object that relative complexity is a somewhat subjective idea since it depends on a choice of simple operations. But in a qubit context a simple operation is one that involves a small number of qubits. For example we might allow all one and two-qubit gates as simple operations. Relative complexity would then be fairly well defined¹ But then a skeptic may say: Yes, but what happens to your definition of complexity if I decide to allow three qubit gates?

I expect (but cannot prove) that the answer is, not much happens at least for large complexity. There are two kinds of ambiguities to worry about—additive and multiplicative. The additive ambiguities are associated with precision—how close do we have to get to the target to declare success? If we are required to get within inner-product distance ϵ , then there is an additive term in the complexity, logarithmic in ϵ . It is closely related to the logarithmic ambiguity in ?? and as we will see in the third lecture, to the $\log \epsilon$ that appears in classical entropy.

There is also a multiplicative ambiguity when we go from allowing only one and two qubit gates to allowing three qubit gates. That is because a three qubit gate can be approximated by some number of one and two qubit gates. That ambiguity can be accounted for as long as the complexity is not too small.

A basic assumption is that there a set of definitions of quantum complexity that exhibit a universal behavior up to additive and multiplicative ambiguities. The important rule is that the allowable gates must be restricted to be k -local with k much smaller than the

¹There are still details associated with how precisely a product of gates must approximate a target state.

number of qubits comprising the system. All k -local means is that the gate involves no more than k qubits.

We'll have more to say about relative complexity but let's first discuss the dual role of unitary operators.

4 Dual Role of Unitaries

Complexity can be defined for states², but in many ways it is more natural to start with the complexity of operations that you can do on a system. That means the complexity of unitary operators. Unitary matrices have a dual role in quantum information theory. First of all they are the matrix representation of unitary operators,

$$U = \sum U_{ij} |i\rangle \langle j| \quad (4.9)$$

Secondly they can be used to represent maximally entangled states of $2K$ qubits,

$$|\Psi\rangle = \sum U_{ij} |i\rangle |\bar{j}\rangle \quad (4.10)$$

In asking about the complexity of unitary operators we are also asking about the complexity of maximally entangled states³,

The space of (special) unitary operators is $SU(2^K)$. It is much bigger than the space of states $CP(2^K - 1)$. Let's begin with a crude estimate of its size.

A unitary matrix in $SU(N)$ has $(N^2 - 1)$ real parameters. If each parameter can take on m values the number of unitary operators is

$$\#unitaries = m^{(N^2-1)}$$

Setting $N = 2^K$,

$$\#unitaries = m^{(4^K-1)}$$

$$\log(\#unitaries) \approx 4^K \log m \quad (4.11)$$

²We continue to work with systems of K qubits. The space of states is 2^K dimensional.

³In the case of maximally entangled systems I mean a restricted form of complexity in which gates are only permitted to act on each side of the entangled system separately, thus preserving the maximal entanglement.

This is the same pattern that we saw for states except that 2^K is replaced by 4^K .

5 Volume of $SU(2^K)$

Let's make a more refined calculation of the number of operators in $SU(N)$ by dividing its volume by the volume of an epsilon ball of the same dimensionality (the dimension of $SU(N)$ is $N^2 - 1$). The volume of $SU(N)$ is (see arXiv:math-ph/0210033)

$$V[SU(N)] = \frac{2\pi^{\frac{(N+2)(N-1)}{2}}}{1!2!3!\dots(N-1)!} \quad (5.12)$$

The volume of an epsilon-ball of dimension $N^2 - 1$ is

$$\frac{\pi^{\frac{N^2-1}{2}}}{\left(\frac{N^2-1}{2}\right)!}$$

Using Stirling's formula, and identifying the number of unitary operators with the number of epsilon-balls in $SU(N)$

$$\begin{aligned} \#unitaries &\approx \left(\frac{N}{\epsilon^2}\right)^{\frac{N^2}{2}} \\ &= \left(\frac{2^K}{\epsilon^2}\right)^{4^K/2} \end{aligned} \quad (5.13)$$

Taking the logarithm,

$$\log(\#unitaries) \approx \frac{4^K}{2} K \log 2 + 4^K \log \frac{1}{\epsilon} \quad (5.14)$$

which is comparable to ???. Again, we see the strong exponential dependence on K and the weak logarithmic dependence on ϵ . The $\log \frac{1}{\epsilon}$ term is multiplied by the dimension of the space.

6 Exploring $SU(2^K)$

We've seen that the space unitary operators is gigantic. Now I want to discuss how to move through it. I've already hinted that we don't make big complexity jumps, but instead

move in little steps called gates. A sequence of gates is called a circuit although it has nothing to do with periodicity. It's just a name.

Definition:

A k -local gate is a nontrivial k -qubit unitary operator chosen from some allowed universal gate set. We assume that if g is in allowed set, so is g^\dagger .

Figure ?? is a schematic representation of a 2-qubit gate.



Figure 1: Time flows from left to right. A gate acts on an incoming state of two-qubits to give an outgoing state.

Gates can be assembled into quantum circuits in the manner shown in figure ??.

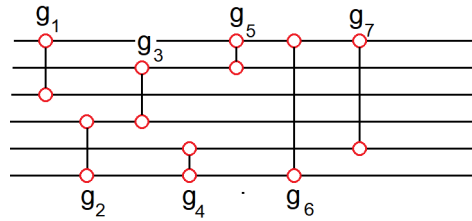


Figure 2: Circuit

A k -local circuit is one made of k -local gates. k -locality is not the same as spatial locality although spatially local circuits are special cases of k -local circuits. In a spatially local circuit the qubits are arranged on a spatial lattice of some dimensionality. Gates are permitted only between near neighbors on the lattice. An example of a k -local and spatially local circuit is shown in figure ??.

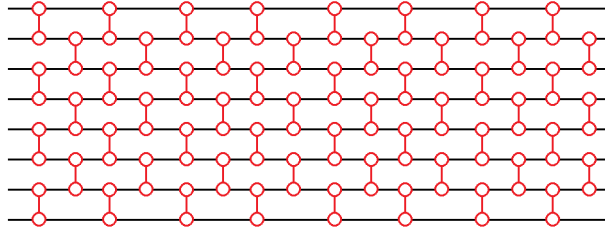


Figure 3: Circuit

Such circuits would be valuable for simulating condensed matter systems but they are not the kind of circuits we will be interested in.

Definition A k -local *all-to-all* circuit is one which is k -local but permits any group of k qubits to interact. Figure ?? is 2-local and all-to-all.

In figure ?? the gates act in series, one gate at a time, but it is more efficient to allow parallel action of gates. The standard k -local all-to-all circuit architecture is shown in figure ??.

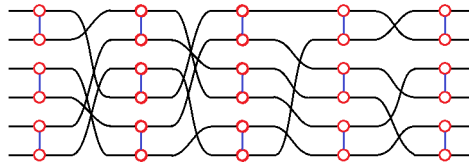


Figure 4: Standard circuit architecture. Note that in each step the gates all commute because they act on non-overlapping qubits.

The rule is that at each step the qubits are grouped into pairs and each pair interacts by means of a two-qubit gate.

We imagine that such a circuit is equipped with a clock, and with each tick of the clock $K/2$ gates act. We'll call this a "step." The number of steps in a circuit is called its depth D . The depth of the circuit in figure ?? is $D = 5$. The number of gates in the circuit is $\frac{KD}{2}$.

Between steps the qubits may be permuted so that any pair can potentially interact. A circuit of this type can be called k -local (2-local in figure ??) and all-to-all. The meaning of

all-to-all is that any pair of qubits can potentially interact. Note that in a given time-step the different gates commute because they act on non-overlapping qubit pairs .

If we measure time τ in units of steps then the number of gates that act per unit time is $K/2$.

A given circuit prepares a particular unitary operator U . Preparing U by a series of steps can be viewed as a discrete motion of a fictitious classical particle (called the auxilliary system in Brown-Susskind arXiv:1701.01107 [hep-th]) on $SU(N)$. The particle starts at the identity operator I and ends at U .

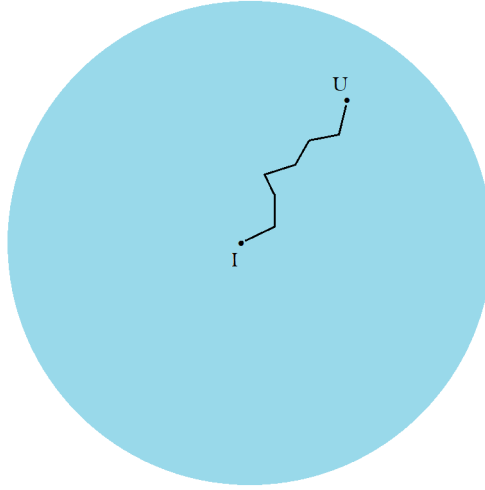


Figure 5

6.1 Relative Complexity of Unitaries

The standard inner product metric for unitaries is similar to the inner product metric for states. In fact if we think of unitaries as the wave functions of maximally entangled systems then it is the inner product for such states. The inner product distance between U and V is,

$$d(UV) = \arccos |\mathbf{Tr} U^\dagger V| \quad (6.15)$$

where \mathbf{Tr} means the normalized trace defined so that $\mathbf{Tr} I = 1$. For the same reasons that I discussed earlier, the inner product distance is not a useful measure of how difficult it is

to go from U to V by small steps. A much better measure is relative complexity.

Given two unitaries the relative complexity $\mathcal{C}(U, V)$ is defined as the minimum number of gates (in the allowed gate set) satisfying,

$$U = g_n g_{n-1} \dots g_1 V \quad (6.16)$$

to within tolerance ϵ . The relative complexity of U and the identity may be defined to be the complexity of U .

$$\mathcal{C}(U) \equiv \mathcal{C}(U, I). \quad (6.17)$$

Figure ?? shows how the relative complexity can be thought of in terms of a discrete curve from V to U .

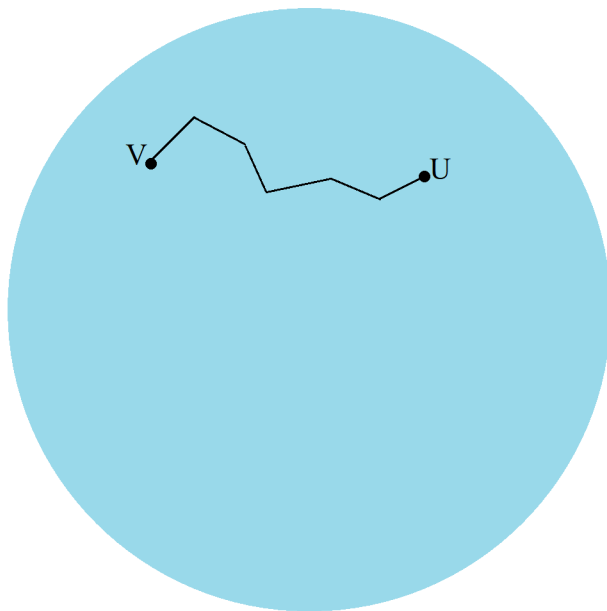


Figure 6: relative-C

Because the curve defining $\mathcal{C}(U, V)$ is the shortest such path, we can think of it as a geodesic, BUT not a geodesic with respect to the inner product metric. It is a geodesic

with respect to relative complexity. The geometry defined by relative complexity is very different from the geometry defined by the inner product distance.

Relative complexity is a metric

Relative complexity satisfies the four defining conditions to be a metric.

1. $\mathcal{C} \geq 0$.
2. $\mathcal{C}(U, V) = 0$ iff $U = V$
3. $\mathcal{C}(U, V) = \mathcal{C}(V, U)$
4. $\mathcal{C}(U, V) \leq \mathcal{C}(V, W) + \mathcal{C}(W, V)$ (Triangle Inequality)

In fact it is a particular kind of metric called *right-invariant*.

Suppose that

$$U = (g_n g_{n-1} \dots g_1) V \quad (6.18)$$

Then for any W it follows that,

$$UW = (g_n g_{n-1} \dots g_1) VW. \quad (6.19)$$

In other words the relative complexity of U and V is the same as that of UW and VW .

$$\mathcal{C}(U, V) = \mathcal{C}(UW, VW) \quad (\text{all } W) \quad (6.20)$$

Thus \mathcal{C} is right-invariant.

On the other hand if we multiply from the left,

$$WU = (W g_n g_{n-1} \dots g_1 W^\dagger) WV. \quad (6.21)$$

But $(W g_n g_{n-1} \dots g_1 W^\dagger)$ generally is not a product of n allowed gates. Therefore \mathcal{C} is not left-invariant⁴.

⁴The usual inner-product metric is both left and right invariant. It is called bi-invariant.

It should be obvious from these remarks that quantum complexity is really a branch of geometry—right-invariant geometry⁵. We’ll build on this idea and discuss two ways of thinking about the geometry of complexity, one in this lecture in terms of graph theory, and the other in lecture 3 in terms of Riemannian geometry.

6.2 Complexity is Discontinuous

In describing relative complexity I’ve suppressed issues having to do with coarse-graining by epsilon balls. For example in ??, ??, ??, and ?? I ignored the phrase, *to within a tolerance epsilon*. In fact as epsilon becomes small it takes an ever increasing number of gates to achieve that tolerance. The increase is only logarithmic in $\frac{1}{\epsilon}$ but nevertheless, with the present definition there is not a formal limit of relative complexity, or of the geometry of relative complexity. One can only define a sequence of geometries as ϵ decreases.

The relation between the familiar inner-product metric and the relative complexity metric becomes wildly discontinuous as $\epsilon \rightarrow 0$. For example two points can be close in complexity space, e.g., $|\mathbf{on}\rangle$ and $|\mathbf{off}\rangle$, and be maximally distant in the inner product metric. Similarly two states can be close in inner product and far in complexity. When this happens the two states have all expectation values close to one another, which means they yield almost identical results for any experiment. Nevertheless making a transition between them requires many gates.

It is an interesting question why a physicist would ever be interested in distinctions which have essentially no effect on expectation values. The answer is that if it were not for black holes, most likely no physicist would be.

Graph Theory Perspective

Circuits can be usefully described using graph theory. I’ll call a graph describing a circuit a *circuit graph*.

Let’s begin at the identity and act with a circuit of depth one, in other words a circuit of a single step with $K/2$ gates. For simplicity let’s assume the allowed gate set is a single non-symmetric two-qubit gate. A choice must be made of how the K qubits are paired. Each pairing will lead to a different unitary. I’ll call the number of possible choices d (an

⁵See Dowling, Nielsen.

Most of the mathematical literature on the subject is about left-invariant geometry but of course this is just a matter of convention.

odd choice of notation but it corresponds with standard graph-theory notation). It's an easy combinatoric exercise to compute d ,

$$d \sim \frac{K!}{(K/2)!} \sim \left(\frac{2K}{e}\right)^{\frac{K}{2}} \quad (6.22)$$

Let's visualize this by a decision tree embedded in $SU(2^K)$. At the central vertex is the identity operator. The branches (or edges) correspond to the d choices.

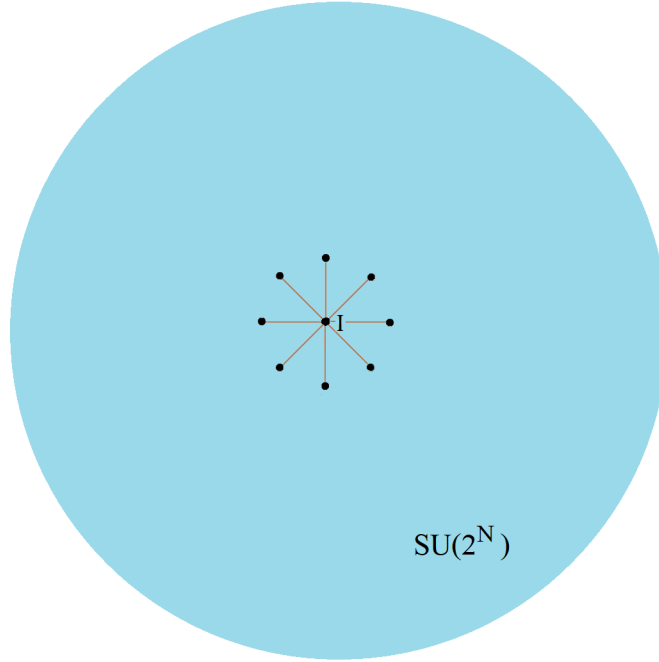


Figure 7: Vertex with $d = 8$.

The number of branches or edges d at a vertex is called the degree of that vertex. In our construction every internal vertex has degree $\frac{K!}{(K/2)!}$. A graph in which every vertex has degree d is called d -regular. But so far the circuit graph is not d -regular because the outer boundary vertices have a single edge. The leaves of the graph are unitary operators in $SU(2^K)$.

Now let's grow the circuit depth by adding another step. I'm going to make one restriction of a technical nature. For large K it makes no difference but it simplifies things. I will assume that when I add a step the choice of pairings is not the same as in

the previous step. This implies that the next layer of the tree has

$$\frac{K!}{(K/2)!} - 1$$

new branches. This is shown in figure ??

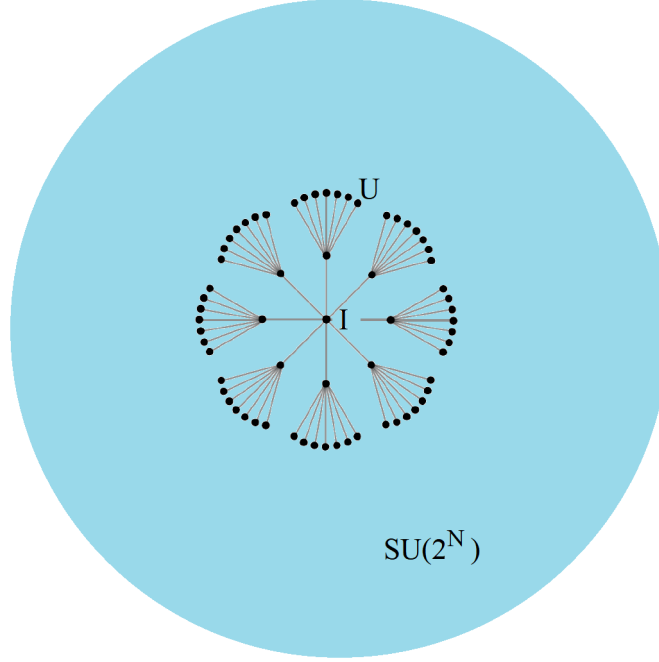


Figure 8: Regular tree with degree 8, depth 2.

Now comes an important assumption which underlies everything I will say. Each endpoint (leaf) of the tree represents a unitary operator. I will assume that the probability of two leaves being the same unitary is vanishingly small. Basically the reason is the extremely high dimension of the space of unitaries. If two leaves represent the same unitary to within ϵ I will say that a collision occurred. My provisional assumption is that collisions are so rare that they can be ignored.

This no-collision assumption must eventually break down as the tree grows outward and we will calculate when that happens. But for now let's ignore collisions.

Assuming no-collisions, the number of unitaries that have been reached at depth D is,

$$\#unitaries = d^D \approx \left(\frac{2K}{e}\right)^{\frac{DK}{2}} \quad (6.23)$$

The number of gates in a circuit of depth D is $DK/2$. Assuming no-collisions, the path to each leaf is the minimal path, implying that the number of gates is the complexity. Thus we can write,

$$\begin{aligned} \#unitaries &= \left(\frac{2K}{e}\right)^{\mathcal{C}} \\ &\sim e^{\mathcal{C} \log K} \end{aligned} \quad (6.24)$$

Let me rephrase this formula in the following way. The sub-volume of $SU(2^K)$ that corresponds to unitaries of complexity \mathcal{C} grows exponentially with \mathcal{C} . It says a number of things. First of all if we think of complexity as the distance from the origin (the unit operator) then it says that the volume grows exponentially with radius—a characteristic of hyperbolic spaces of negative curvature. This may seem surprising: with the usual bi-invariant metric $SU(2^K)$ is positively curved. But we are talking about a different metric, relative complexity. This negative curvature is a symptom of chaos and in my opinion it, not fast scrambling, is the general signature of quantum chaos.

Equation ?? suggests something else. In classical statistical mechanics the volume of phase space associated with states of a given entropy is exponential in the entropy. The exponential growth of volume in $SU(2^K)$ associated with a given complexity is the basis for a deep analogy between complexity and entropy including a *Second Law of Complexity*. It can be summarized by the slogan: *Complexity is the entropy of the auxiliary system \mathcal{A} .*

One more point about the circuit graph and its embedding in $SU(2^K)$. Figures ?? and ?? are very schematic. It is difficult to accurately convey the actual properties of the embedding with pictures given the fact that the space is extremely high dimensional. One thing to keep in mind is that with the usual bi-invariant metric, $SU(2^K)$ is small in the sense that the largest distance is $\pi/2$. A single one-step circuit will move U about that distance. So although the tree abstractly grows outward from the center as in the left panel of figure ??, the embedding in $SU(2^K)$ looks more like the right panel.

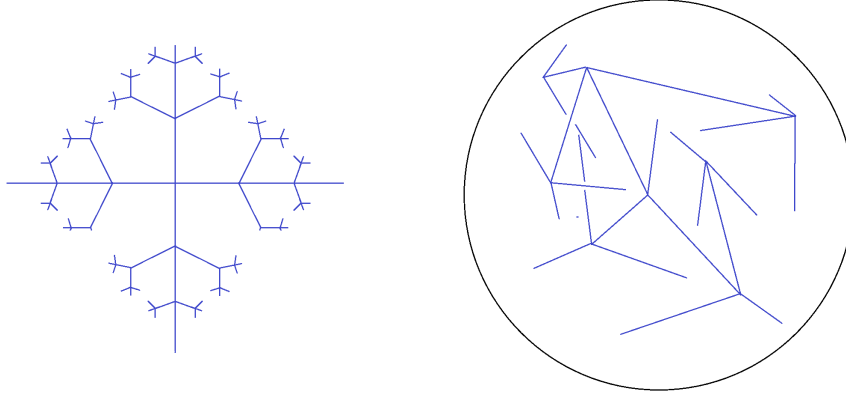


Figure 9: Two views of the circuit tree. On the left is an abstract decision tree. The complexity of the unitary leaf grows linearly with the number of steps from the identity at the center. Relative complexity is proportional to the number of steps between leaves. On the right the tree is embedded in $SU(2^K)$ with the usual inner product metric.

The epsilon balls get filled in by the growing circuit in a very fractal manner which is related to the fact mentioned earlier that complexity is discontinuous as $\epsilon \rightarrow 0$.

6.3 Collisions and Loops

Each vertex of the tree represents a unitary operator and for that reason we can think of the tree as being embedded in $SU(2^K)$. Suppose a collision does occur. This means that two leaves of the tree are located at the same point in $SU(2^K)$. We can represent this by drawing the two leaves as a single leaf as in figure ??.

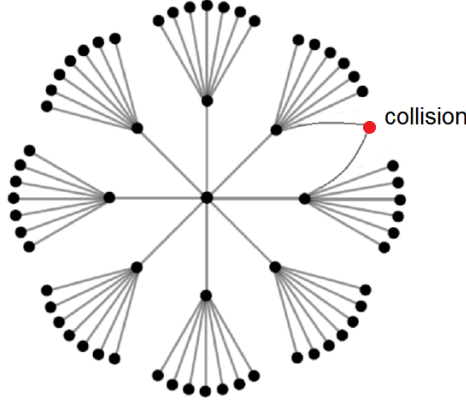


Figure 10

The figure illustrates the fact that collisions induce loops in the graph. This allows us to translate the rarity of collisions into graph-theoretic terms. No collisions at all would mean no loops, i.e., exact tree-ness. But as we will see, eventually for very large depth, collisions must happen. This means that very large loops must occur. The correct formulation, which we will come back to, is that small loops must be absent or very rare.

Since the number of epsilon-regulated unitaries is finite we will eventually run out of room on $SU(2^K)$. That happens when the number of leaves (as given by ??) is equal to the total number of epsilon-balls as given in ??. This determines the maximum possible complexity.

$$\left(\frac{2K}{e}\right)^{C_{max}} = \left(\frac{2^K}{\epsilon^2}\right)^{4^K/2} \quad (6.25)$$

or

$$C_{max} = 4^K \left[\frac{1}{2} + \frac{|\log \epsilon|}{\log K} \right] \quad (6.26)$$

Again, strong dependence on K , weak dependence on ϵ . Roughly

$$C_{max} \sim 4^K.$$

Apart from a factor of K which is swamped by the exponential, this is also the depth at which collisions must occur. In other words it is the maximum radius at which the tree stops being tree-like. Finally it is also the largest distance between nodes of the tree—the diameter of the graph.

We can now state the no-collision assumption more precisely.

Loops smaller than 4^K are absent or very rare.

In graph-theoretic terms, the girth of the graph is $\sim 4^K$.

Another point follows from the fact that the total number of unitaries is $\sim e^{4^K}$. We may identify this with the number of vertices in the graph. This implies that the diameter of the graph is logarithmic in the number of vertices.

The breakdown of no-collisions

Let's consider what happens when the tree-ness breaks down. Up to that point the graph is a d -regular tree similar to figure ?? except with much higher degree.

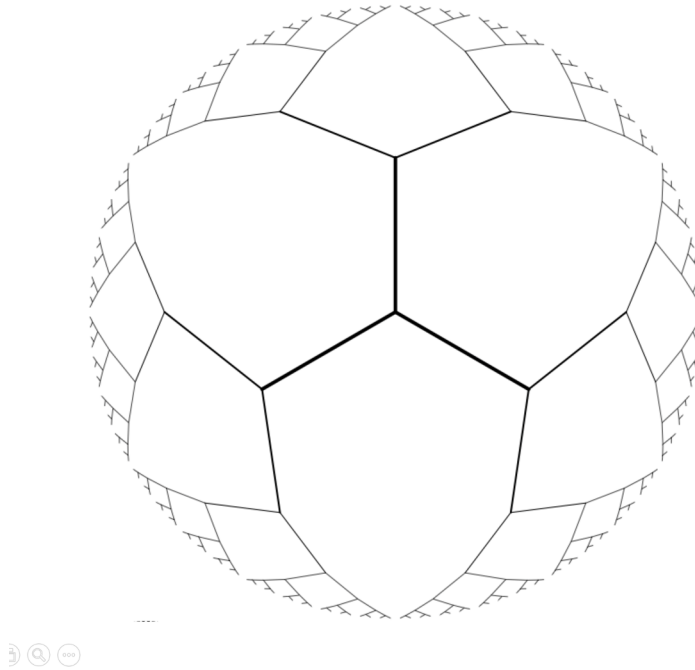


Figure 11: Regular tree with degree 3

But once we reach \mathcal{C}_{max} the graph can not continue to grow. Collisions occur and loops must form. The graph must double back on itself and revisit previously visited epsilon-balls⁶. We show a couple of possible loops that might form in figure ??.

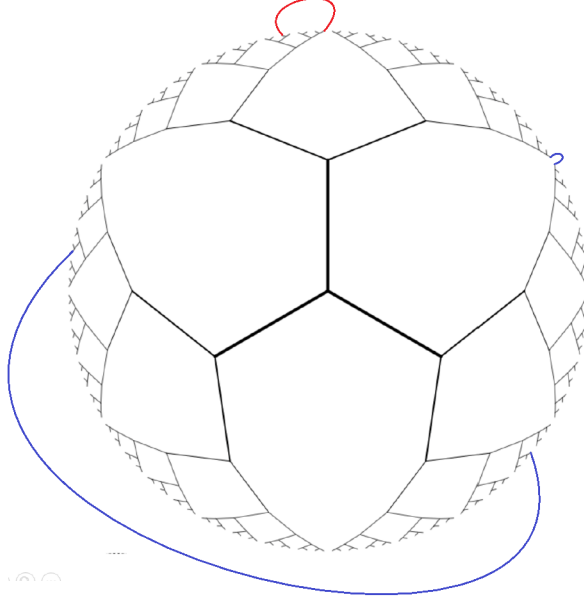


Figure 12: Graph ceases being a tree and doubles back. Since group is homogeneous the structure must look the same from every vertex. The shortest loops (girth of graph) is of order 4^K . Thus the red loop is too short. The loops must look more like the blue loop.

Now we know that any loop which passes through the central node must be very big, namely of length $\sim 4^K$. But because the space $SU(2^K)$ is a group space, every point on the graph is the same as every other point. Thus it must be that loops passing through any point must be equally long. It is clear from the figure that loop containing the red segment is much shorter and should not occur, but the loop containing the blue segment is long and may occur.

Let me summarize the conjectured properties of circuit graphs generated by iterating one-step circuits:

⁶If we follow an epsilon ball from the identity it will not in general perfectly coincide with an epsilon ball after executing a loop. There is a bit of sloppiness but it doesn't seem to be important. Note that if ϵ is decreased the maximum complexity increases and the graph becomes bigger.

1. The degree is the same for all vertices and is given by

$$\text{degree} = \frac{K!}{(K/2)!} \sim \left(\frac{2K}{e} \right)^{\frac{K}{2}}. \quad (6.27)$$

This is much smaller than the number of vertices. The graph can be said to be sparse.

2. The number of vertices in the graph is of order e^{4K} .
3. The greatest distance between vertices (diameter) is $\mathcal{C}_{max} \sim 4^K$. The diameter is therefore logarithmic in the number of vertices.
4. Loops of length less than 4^K are rare or absent.
5. The graph is homogenous and from any point looks tree-like out to distances of order the diameter.

These properties are very familiar to graph theorists. They are the properties of a good *expander* graph. I think it may be even stronger. It may be among the strongest expanders called Ramanujan graphs. Graphs of this type are discrete analogs of finite negatively curved geometries such as the hyperbolic plane, with identifications that render it compact⁷

Quantum complexity has a strong geometrical aspect that was first appreciated by Nielsen. I would make the case that it is basically the subject of right-invariant geometries on a group space with the parameters chosen so that the curvature is negative. In the third lecture I will come back to this view of complexity theory.

⁷The relation between complexity and such geometries was described in Brown, Susskind and Zhao, [arXiv:1608.02612 [hep-th]].

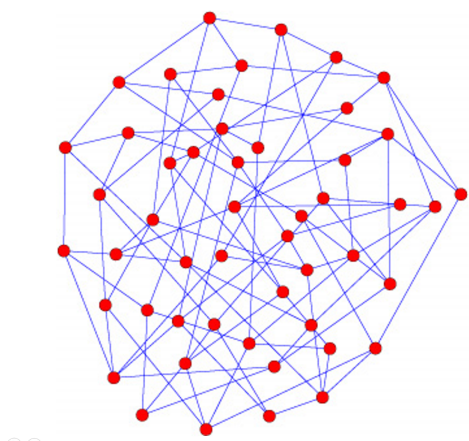


Figure 13: 50-vertex, (almost) regular $d = 4$ expander. Generated by some AI optimization protocol.

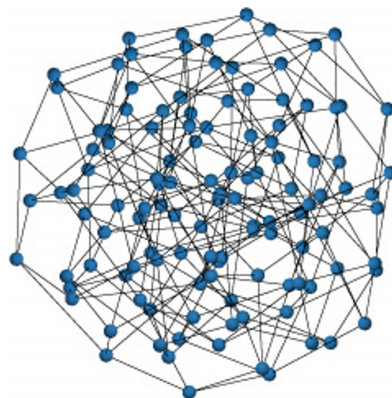


Figure 14: 120 vertex, regular $d = 4$ Ramanujan-expander. Quoting Donetti, et al.: “Some remarks on its topological properties are in order. The average distance 3.714 is relatively small (i.e. one reaches any of the 120 nodes starting from any arbitrary origin in less than four steps, on average.” Note: $\log 120 = 4.78749174278$. Not too different from 3.714. The authors then go on to say, “The clustering coefficient vanishes, reflecting the absence or short loops (triangles), and the minimum loop size is large, equal to 6, and identical for all the nodes. In a nutshell: the network homogeneity is remarkable; all nodes look alike, forming rather intricate, decentralized structure, with delta-peak distributed topological properties.”

7 The Second Law of Quantum Complexity

I'm not sure why complexity theorists never remarked on the similarity of quantum complexity and classical entropy, or the existence of a second law of complexity⁸. The second law is the subject of lecture three but I will briefly explain it here because it plays an important role in lecture two.

Just to be clear, although ensemble-averaged complexity is a kind of entropy it is *not* conventional entropy. The conventional entropy of a system of K qubits is bounded by $K \log 2$, the logarithm of the maximum number of mutually orthogonal vectors in the Hilbert space. By contrast the maximum complexity is exponential in K . Maximum complexity $\sim 4^K \log 1/\epsilon$, the number of ϵ -balls in $SU(2^K)$.

The quantity $4^K \log 1/\epsilon$ does have an interpretation in terms of the entropy of a classical *auxiliary* system \mathcal{A} associated with the quantum system of K qubits. We may think of it as the maximum entropy of a classical system with 4^K classical degrees of freedom. The auxiliary system is just the classical 4^K collection of coordinates⁹ that describe the evolving time-evolution operator $U(t)$. Figure ?? illustrates the auxiliary system. In lecture three this will be made more precise. Quantities describing the auxiliary system will carry a subscript \mathcal{A} .

For now let us consider a simplified version of the quantum evolution of a system at high temperature. We envision an ensemble of fictitious particles moving on $SU(2^K)$. The particles are random walkers which all start at the origin of $SU(2^K)$ (the identity operator) at some initial time. The dynamics is discrete: at each step the position is updated by applying a depth-one circuit. This means the particles execute random walks on the graph that I just explained.

At each vertex the decision for the next step is made randomly. I will allow the possibility of back-steps along the previous edge. Initially the probability is concentrated at the origin and the entropy of the fictitious system is zero.

After one time-step the particle is at the first level of the tree on one of the leaves, as in figure ?. The fictitious entropy is $\log d$ and the complexity is $\frac{K}{2}$. In the next step the particle has a probability $\sim 1/d$ to back-track, but for large d that is negligible. With

⁸There is a very beautiful paper by Zurek in which he argues that the classical entropy of a system is the ensemble average of the Kolmogorov algorithmic complexity of states in phase space. [W. Zurek, Phys. Rev. A **40**, 4731 (1989)] The idea is very similar to the connection between quantum-computational-complexity and entropy described by Brown et.al in [arXiv:1701.01107 [hep-th]], and described here.

⁹Strictly speaking, $4^K - 1$.

probability close to one the particle moves outward to the next level where the complexity is K . and the auxiliary entropy $S_{\mathcal{A}}$ is $\log d^2 = 2 \log d$.

After n steps the particle with high probability is at the n^{th} level, the auxiliary entropy is $S_{\mathcal{A}} = n \log d$, and the complexity is $nK/2$. Evidently the complexity and fictitious entropy of the auxiliary system are related,

$$S_{\mathcal{A}} \approx \mathcal{C} \log K \quad (7.28)$$

This identification is dependent on the negative curvature and high dimensionality of complexity space. These two ingredients are what insure that collisions are rare and that we can identify the level of the tree with the minimum distance from the origin at I . In other words, up to a factor $K/2$ we may identify the depth with complexity.

Of course it is not rigorously true that there are no collisions. It's just that collisions are rare for sub-exponential time. We can be reasonably sure that almost all leaves (vertices) have complexity proportional to their level, but it is much harder to know that a given leaf has not had collisions in the past. It's for this reason that we identify the entropy $S_{\mathcal{A}}$ with the ensemble averaged complexity.

The second law of complexity is just the second law of thermodynamics—the overwhelming statistical likelihood that entropy will increase—applied to the ensemble average of complexity. The reason why complexity almost always increases when it is less than maximum is the same as why classical entropy almost always increases when it is less than maximum—the number of states exponentially increases with increasing entropy/complexity.

Let us follow a particular member of the ensemble. As long as it is not an exponential number of steps from I the complexity will simply reflect the exponentially growing number of states as we move outward from the origin. It will with very high probability increase linearly with time. However once $\mathcal{C} \sim 4^K$ the particle will have reached the maximum distance on the graph and the complexity will stop increasing. Complexity equilibrium will have been achieved. The number of states with maximum complexity is so vast, that the particles will get lost among them and remain at maximum complexity for a recurrence time. The recurrence time for the classical system will be $t_{recur} = \exp S_{\mathcal{A}}$ which is doubly exponential in K .

$$t_{recur} \sim e^{4^K} \quad (7.29)$$

Thus we expect a singly exponential time $\sim 4^K$ during which complexity linearly grows, after which it remains approximately constant at its maximum. But then, on gigantically long time scales $\sim \exp \exp K$ it will recur to small values, and $U(t)$ will return to the neighborhood of the identity.

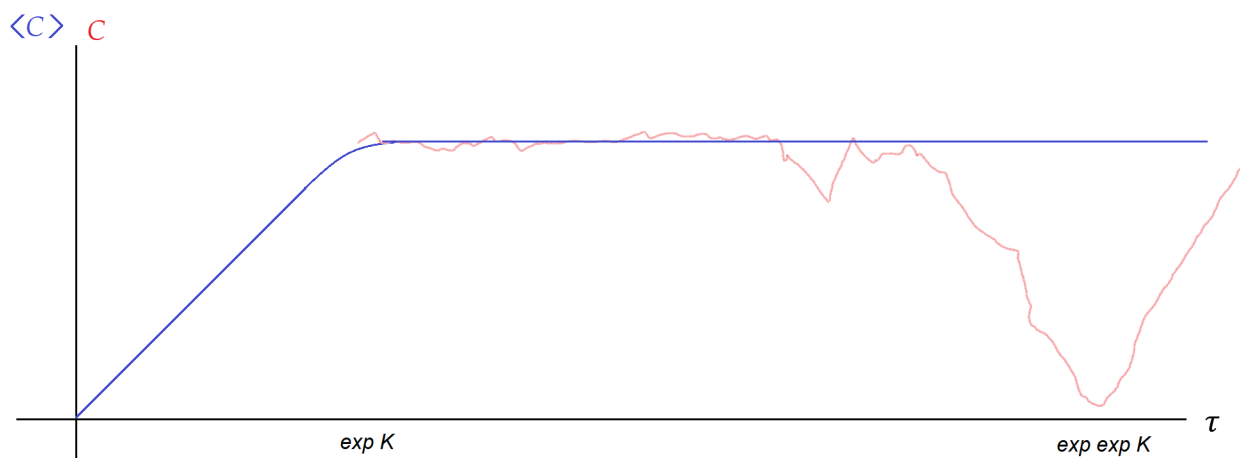


Figure 15

The transition from linear growth to complexity equilibrium is very sudden.

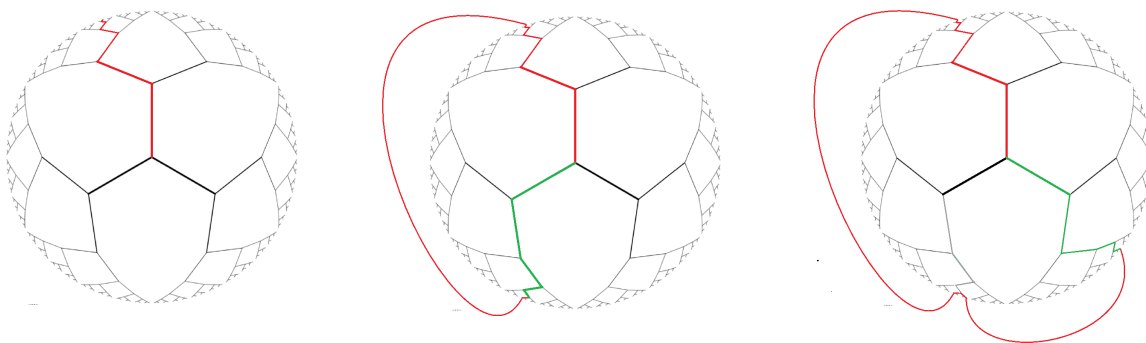


Figure 16

It is interesting to see how this behavior can be understood in terms of circuit graphs.

We can get an idea by looking at figure ?? . The figure shows the evolution of the \mathcal{A} auxiliary system moving according to some “dynamical” rule that mocks up Hamiltonian evolution starting with $U(0) = I$. The trajectory moves outward and until collisions occur the complexity increases linearly with time. The complexity is proportional to the graph distance from I . That’s shown as the red trajectory in the left panel.

Once the trajectory reaches maximum complexity it keeps going, but it has no choice but to visit previously visited sites. With overwhelming probability it will jump to another almost maximally complex state. That is shown in the middle panel.

But now, there is a shorter path to the end of the red trajectory. It is shown in green. The complexity is **NOT** the length of the red trajectory but rather the length of the green trajectory. Thus the complexity does not increase and may even decrease a bit.

From there the dynamical red trajectory continues. With overwhelming likelihood it moves outward because the overwhelming number of branches reach outward. It soon reaches the next point where it has to loop around. A new green trajectory forms that is quite different than the previous one, but also has close to maximal complexity. That is the reason why the top of the curve in figure ?? is a bit ragged.

Eventually the auxiliary particle will find its way back to low complexity and the cycle will repeat but this takes a quantum-recurrence time.

Note that the transition from linearly increasing complexity to equilibrium at the top of the curve is sudden. It’s similar to a first order phase transition where another local minimum crosses over and becomes the global minimum.